

UNCLASSIFIED

AD NUMBER	
AD335934	
CLASSIFICATION CHANGES	
TO:	unclassified
FROM:	confidential
LIMITATION CHANGES	
TO:	Approved for public release, distribution unlimited
FROM:	Distribution authorized to U.S. Gov't. agencies and their contractors; Administrative/Operational Use; MAR 1963. Other requests shall be referred to Office of Naval Research, Arlington, VA 22203-1995.
AUTHORITY	
31 Mar 1975, DoDd 5200.10; ONR ltr, 15 Jun 1977	

THIS PAGE IS UNCLASSIFIED

THIS REPORT HAS BEEN DELIMITED
AND CLEARED FOR PUBLIC RELEASE
UNDER DOD DIRECTIVE 5200.20 AND
NO RESTRICTIONS ARE IMPOSED UPON
ITS USE AND DISCLOSURE.

DISTRIBUTION STATEMENT A

APPROVED FOR PUBLIC RELEASE;
DISTRIBUTION UNLIMITED.

CONFIDENTIAL

335 934

*Reproduced
by the*

DEFENSE DOCUMENTATION CENTER

FOR

SCIENTIFIC AND TECHNICAL INFORMATION

CAMERON STATION, ALEXANDRIA, VIRGINIA



NOTICE: When government or other drawings, specifications or other data are used for any purpose other than in connection with a definitely related government procurement operation, the U. S. Government thereby incurs no responsibility, nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto.

NOTICE:

THIS DOCUMENT CONTAINS INFORMATION
AFFECTING THE NATIONAL DEFENSE OF
THE UNITED STATES WITHIN THE MEAN-
ING OF THE ESPIONAGE LAWS, TITLE 18,
U.S.C., SECTIONS 793 and 794. THE
TRANSMISSION OR THE REVELATION OF
ITS CONTENTS IN ANY MANNER TO AN
UNAUTHORIZED PERSON IS PROHIBITED
BY LAW.

CONFIDENTIAL

Copy No. *111* of *175*
81613

335934

ASTIA

CATALOGED BY

ADU HQ.

2

HUMAN FACTOR PROBLEMS IN ANTI-SUBMARINE WARFARE

Technical Report 206-21

**THE DISPLAY OF PROBABILISTIC SOLUTIONS
IN SONAR TARGET CLASSIFICATION (U)**

335 934

DOWNGRADED AT 3 YEAR INTERVALS
DECLASSIFIED AFTER 12 YEARS
DOD DIR 5200.10

ASTIA
MAY 2 1963
RECEIVED
TISIA A

HUMAN FACTORS RESEARCH, INCORPORATED
1112 Crenshaw Boulevard • Los Angeles 19, California • WEBster 3-7356

CONFIDENTIAL

CONFIDENTIAL

HUMAN FACTOR PROBLEMS IN ANTI-SUBMARINE WARFARE

Technical Report 206-21

**THE DISPLAY OF PROBABILISTIC SOLUTIONS
IN SONAR TARGET CLASSIFICATION (U)**

Robert R. Mackie

Prepared for

**Personnel and Training Branch
Psychological Sciences Division
Office of Naval Research
Department of the Navy**

by

**Human Factors Research, Incorporated
Los Angeles 19, California**

**March 1963
Contract Nonr 2649(00)
NR 153-199**

CONFIDENTIAL

CONFIDENTIAL

This document contains information affecting the national defense of the United States within the meaning of the Espionage Laws, Title 18, U.S.C., Sections 793 and 794, the transmission or revelation of which in any manner to an unauthorized person is prohibited by law.

CONFIDENTIAL

CONFIDENTIAL

Reproduction in whole or in part is
permitted for any purpose of the
United States Government.

iv

CONFIDENTIAL

CONFIDENTIAL

ABSTRACT

Current sonar equipments, as well as other sensor systems employed in ASW, produce only probabilistic evidence of the target's nature. That is, the classifier is always faced with some degree of uncertainty even after all clues have been correlated and evaluated properly. Traditionally the Navy has recognized this fact by officially assigning targets to categorical classes such as "probable submarine," "possible submarine" and "nonsubmarine."

A small number of categorical classifications deprives the decision maker of useful information available in the sonar clues and eventually restricts his alternative courses of action. Theoretically, the number of correct decisions in a probabilistic information system will be maximized if they are based on some monotonic function of likelihood ratio.* It follows that the number of appropriate actions will be maximized as well.

The present investigation was concerned with the comparative meaningfulness of five alternative methods of displaying classification information. Three displays were based on likelihood ratio, a fourth was a display of 20 solution lights such as that used with MITEC, and the fifth was based on the traditional Navy three-category classification (trichotomy). Navy officers and senior petty officers served as subjects.

The three displays based on likelihood ratio produced interpretations very similar to those of a theoretically ideal interpreter. There was some unwillingness to extract the appropriate amount of certainty from very small or very large ratios, a finding in line with those of other investigators.

*Likelihood ratio is here defined as the ratio of the probability that a particular configuration of clues resulted from a submarine target divided by the probability that it resulted from a non-submarine target.

CONFIDENTIAL

CONFIDENTIAL

The 20-Lights display was interpreted essentially as a linear, equal-interval scale of probability, an interpretation not intended by its designers. Overly conservative reactions, in relation to an ideal interpreter, occurred both for moderately low and moderately high probabilities.

The traditional Navy trichotomy produced severely limited interpretations, the phrase "possible submarine" producing reactions of nearly maximum uncertainty. There were substantial differences in interpretation among the subjects.

The problem of intelligent design of probabilistic displays is discussed along with the need of the designer of clue correlating devices to match the output display with the interpretations of potential users. The usefulness of expressions of certainty and the amount of variance in these expressions as intermediate criteria of the meaningfulness of the displays is pointed out.

CONFIDENTIAL

CONFIDENTIAL

ACKNOWLEDGEMENTS

The author is indebted to Captain L. P. Spear, Director, and Lieutenant Commander G. W. Bird of Tactical Training, ASW School, San Diego, and to Commander R. Plunkett, Training Officer, for their permission to conduct this study and willingness as participants.

CONFIDENTIAL

CONFIDENTIAL

Table of Contents

	Page
ABSTRACT	v
ACKNOWLEDGEMENTS	vii
LIST OF TABLES AND FIGURES	xi
I. THE GENERAL PROBLEM	1
THE SPECIFIC SETTING	1
Correlation of Clues	1
Scale of Resemblance	3
The Traditional Navy Trichotomy	5
The HHIP Dichotomy	7
The MITEC 20 Final Lights	10
THE EXPERIMENTAL QUESTION	12
II. PROCEDURE	15
Hypothetical Distributions of Submarine and Nonsubmarine Targets	15
Selection of Observation Points on Scale of Resemblance	16
Alternative Methods of Display	17
The Response Scale	20
The Subjects	22
III. RESULTS	
Expressions of Certainty	25
Differences in Interpretation Between Subjects	32
Differences in Interpretation Between Officers and Petty Officers	34

CONFIDENTIAL

CONFIDENTIAL

Table of Contents (Cont.)

	Page
IV. DISCUSSION	37
Display Design and the Evaluation Function .	37
Display Design and Tactical Flexibility . . .	39
Designing the Displays vs. Programming the Output	40
REFERENCES	43

x

CONFIDENTIAL

CONFIDENTIAL

List of Tables and Figures

Table		Page
I	Estimated Frequencies (f) at 12 Observation Points and Corresponding Likelihood Ratios (l) for Hypothetical S/M and N/S Contact Distributions	17
II	Average Certainty Ratings for Five Displays Related to Theoretically "Ideal" Observer (N=23) . .	27
III	Inter-Judge Agreement as Reflected by Dispersion of Certainty Ratings (σ) (N=23)	33
IV	Differential Confidence as to Target's Nature . .	35
Figure		
1	Overlapping frequency distributions of submarine and nonsubmarine clue configurations on a hypothetical resemblance scale	4
2	Three observation points on resemblance scale with different frequencies of occurrence for submarine and nonsubmarine clue combinations	9
3	Format of 20-lights display used with MITEC MK. I Mod. 1	11
4a	Relative Frequency of occurrence display	18
4b	Bar Graph display	18
4c	Relative Odds display	19
4d	20-Lights display	19
4e	Navy Trichotomy display	20
5	Response scale for indicating level of certainty .	21
6	Certainty ratings produced by three displays of relative likelihood. (Each point based on 23 ratings)	26

CONFIDENTIAL

List of Tables and Figures (Cont.)

Figure		Page
7	Mean certainty and dispersion of ratings produced by the 20-Lights display (23 observers)	29
8	Mean certainty and dispersion of ratings produced by traditional Trichotomy display (23 observers) .	30

CONFIDENTIAL

THE DISPLAY OF PROBABILISTIC SOLUTIONS IN SONAR TARGET CLASSIFICATION (U)

I. THE GENERAL PROBLEM

Short of visual sighting, and in some cases even with visual sighting, the classification of a sonar contact invariably is made with some degree of uncertainty. None of the sensing mechanisms currently available to the ASW forces provides sufficiently well-defined information about the characteristics of a target to enable the classification to be made with complete certainty. Different kinds of sensors, i.e., active sonar, passive sonar, radar, MAD, ECM, differ in the degree to which they provide useful classification information. Further, there are large differences in the reliability with which they provide it.

The present investigation was concerned with the meanings conveyed by several alternative ways of presenting uncertain, or probabilistic target information to the decision maker.¹ The investigation is germane to all types of ASW systems and their respective sensors. While the study was performed in the context of the ASW surface ship employing active sonar, the problem is common to all systems involving decision making in the face of uncertain or incomplete information.

THE SPECIFIC SETTING

Correlation of Clues

No single display in an active sonar system provides an adequate basis for target classification (1, 2). Consequently,

¹The ultimate classification of an underwater target is made by the commanding officer of the investigating vehicle, or, in some cases, by a superior tactical commander.

CONFIDENTIAL

CONFIDENTIAL

classification in these systems typically depends on the correlation of clues from three basic types of display;²

1. An audible presentation of each target echo via headphones or loudspeaker;
2. A visual presentation of each target echo on a scope; and
3. A time-range history of a series of target echoes on a graphic recorder.

Although each of these displays is activated by the same basic target signals, the information they provide to the observer is not wholly redundant. Further, the degree to which the information in a particular display is useful varies considerably from target to target and time to time depending upon a number of environmental and operational variables. These variables are only partially under the control of the operators and their effects are not always predictable.

While each display typically produces imperfect information, it has been found that the proper correlation of clues from the several displays significantly increases the likelihood of achieving a correct target classification (2, 3, 4, 5). The correlation process, while logical, creates heavy demands on memory and reasoning processes. Consequently, the need has been recognized for providing the decision maker with a mechanical aid to the correlation and evaluation of clues.

After a number of early failures, two devices potentially fulfilling this need recently have been built. The first was the HHIP³ currently in fleet use, and the only such device thus far to enjoy substantial success in the fleet. The second was the

² Additional sources of information, such as DRT plot, also affect the classification. The present discussion, however, is oriented toward the information generated solely by sonar displays.

³ Hand-Held Information Processor.

CONFIDENTIAL

CONFIDENTIAL

MITEC⁴ currently in the process of technical and fleet evaluation (5).

The underlying mechanics of these devices is beyond the scope of this report. It has been demonstrated that both have valid, though somewhat different, classification logics. Neither device, however, produces a completely certain solution. The question then arises as to how to display the various levels of solution, or relative certainty of the solutions produced. Unless the display conveys the results in the most meaningful fashion, the most appropriate tactical course of action may not follow.

Scale of Resemblance

It is convenient to think of the great number of various possible combinations of target clues as falling on a scale that reflects the extent to which the target possesses the physical features and behavioral characteristics of the submarine. Theoretically, every configuration of clues from a submarine contact should fall at a high point on this "scale of resemblance." In practice, many do not, the distribution being negatively skewed to a marked degree. On a given occasion the submarine target may fall at any point along this scale (see Figure 1). Further, the frequency with which clue combinations from submarine targets fall very high on this scale, e.g., values 19 or 20, may actually be smaller than the frequency with which they occur at lower scale values, e.g., 16, 17 or 18. The reason, of course, is that the discriminative properties of the sensors (in this case the active sonar) are far from perfect. The vagaries of sound transmission and severe energy loss in the ocean often seriously degrade the quality of the information in the target echo. The alignment and calibration of the receiving and processing equipment also affect the displayed clues.

⁴Modular Integrated Target Echo Classifier.

CONFIDENTIAL

CONFIDENTIAL

The displays are not ideally designed from a classification point of view. And the human operator as a perceiver of the information often performs unreliably.

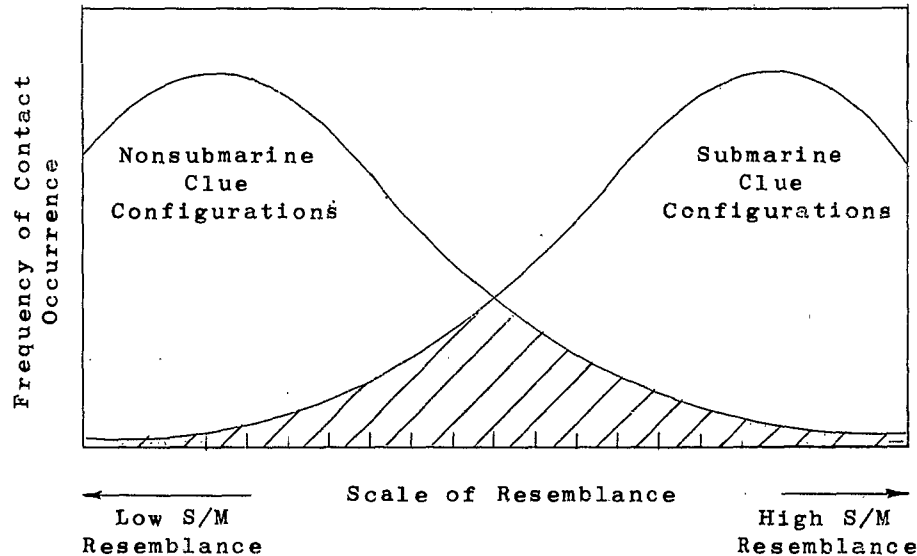


Figure 1. Overlapping frequency distributions of submarine and nonsubmarine clue configurations on a hypothetical resemblance scale.

Configurations of clues from nonsubmarine targets also fall at various points along the scale of resemblance. While a detailed analysis of many nonsubmarine contacts eventually indicates that they are quite unlike a submarine, a substantial number display many submarine characteristics, at least for a period of time.

There may be few clue configurations from nonsubmarine targets that fall at very high points on the scale of resemblance but there will be substantial numbers that occupy mid-scale positions as indicated by the overlap of the curves in Figure 1. It is in this region where the classifier will be maximally uncertain as to the

CONFIDENTIAL

CONFIDENTIAL

nature of his target. Further, complete certainty is not attainable, even at the extreme ends of the scale.

It is clear, then, that the decision maker responsible for the ultimate classification of a sonar contact is faced with a situation having the following characteristics:

1. The contact will present a configuration of clues which, when properly correlated, will place it at some point along a scale of submarine resemblance.
2. In general, high values on this scale imply high likelihood that the target is a member of the submarine population, and conversely,
3. In view of the limitations on present sensors, signal processing techniques and displays, no point on this scale can be regarded as revealing the target's nature with certainty.

Thus, the decision maker is confronted with a probabilistic solution whether he recognizes it or not. The question is, should some attempt be made to display the relative probabilities to him and, if so, in what form?

The Traditional Navy Trichotomy

In a sense, the probabilistic nature of sonar classifications has long been recognized. Official Navy doctrine as promulgated in NWIP 24-1 provides for four levels of classification and sets forth specific criteria for each. These classifications, and the criteria that pertain to sonar, are as follows:

- a. "Positive Submarine." "This classification is given only when a surfaced submarine, submerged submarine, periscope, snorkel, or a torpedo obviously fired by a submarine is sighted and identified by personnel competent to positively recognize such an object." (Note: This classification cannot be given on the basis of sonar evidence.)

CONFIDENTIAL

CONFIDENTIAL

b. "Probable Submarine."

(1) "Sonar Contact. Sonar has picked up a probable submarine when all of the following are observed and reported by competent personnel:

1. "Tracking of a firm contact for a total of at least 15 minutes with courses, speeds, echo quality, and doppler effect compatible with submarine characteristics and confirmed by appropriate change of target aspect in chemical recorder traces and by the target track as plotted on the DRT or by other means. Or, tracking of a medium speed (in excess of 10 knots) firm contact for any period of time with courses, echo quality, turbine whine, and doppler effect compatible with nuclear submarine characteristics and confirmed by appropriate change of target aspect in chemical recorder traces and by the target track as plotted on the DRT or by other means.
2. "One or more of the following are heard or detected--radar contact confirming the sonar track, propeller sounds, internal submarine noises, sonar signals, sonar jamming, starting and securing of snorkel operations, or pronounced wake effect if track indicates submarine moving at sufficient speed."

c. "Possible Submarine." "The possible submarine classification may be given by competent personnel under . . . the following condition(s):

1. "Julie, sonar, radar, ECM, sono-buoy, or MAD contacts are investigated or tracked without confirming all of the characteristics listed for probable submarine, but are suspected to be of submarine origin."

CONFIDENTIAL

CONFIDENTIAL

- d. "Nonsubmarine." "Classification of a contact exhibits or later develops nonsubmarine characteristics or due to other implausible circumstances is adjudged by competent personnel as nonsubmarine."

Insofar as sonar is concerned, only categories (b), (c) and (d) above are legitimate classifications. This three-way grouping is herein referred to as the Navy trichotomy, and while ordinarily given verbally, can be thought of as the way in which sonar target classifications traditionally have been relayed (displayed) to the decision maker.

The first two classifications are probabilistic answers stated in unquantified form. Presumably the third "Nonsubmarine" is a statement of certainty. In practice, however, there undoubtedly are elements of uncertainty about this conclusion and, as indicated before, certainty of any kind is unlikely to be obtained on the basis of sonar information alone. Further, the transition from "Nonsubmarine" to "Possible Submarine," and from "Possible Submarine" to "Probable Submarine," will rarely be abrupt since most targets will meet some but not all of the criteria specified for each category.

The practical meaning of the terms "Probable Submarine," "Possible Submarine" and "Nonsubmarine," in contrast to their doctrinal meaning, has not been the subject of study and is consequently unknown. The possibility exists that quite different meanings are conveyed to different people. A part of the present study therefore was aimed at trying to determine the meaning of these terms in relation to the meanings conveyed by other display alternatives.

The HHIP Dichotomy

The HHIP, based on the NEL system of clue correlations, essentially provides the decision maker with a two-category or dichotomous

CONFIDENTIAL

CONFIDENTIAL

output: submarine or nonsubmarine.

The device also occasionally displays the conclusion "conditional submarine" which occurs when the target displays certain characteristics of the beam aspect submarine. Since it is often difficult for sonar to discriminate the beam submarine from the nonsubmarine domain, the term "conditional submarine" is employed, not as a definitive output, but rather as an instruction to the decision maker to seek further information before a conclusion is reached.

The HHIP has been programmed to maximize the number of correct conclusions obtainable assuming that a dichotomous output is necessary or desirable. The decision to designate a contact as "submarine" or "nonsubmarine" is based on the likelihood ratios resulting from analyses of clue configurations produced by a substantial number of recorded sonar contacts.⁵

If the information sensed by the sonar and displayed to the operator resulted in perfect discrimination, that is, if it produced a 2-point distribution, or even two continuous distributions with a minimum of overlap, the output of the HHIP would be ideal. It effectively removes all possibility of subjective differences in interpretation. Since configurations of submarine and nonsubmarine clues fall into two overlapping frequency distributions, however, instead of into a 2-point distribution, a question legitimately can be raised concerning the failure of the HHIP to differentiate among targets whose probabilities of being a submarine are quite different. For example, it would not distinguish among three targets whose clue configurations have, as shown in Figure 2, quite

⁵ Likelihood ratio in this case is the ratio of the probability that a particular observed configuration of clues resulted from a submarine target divided by the probability that it resulted from a non-submarine target.

CONFIDENTIAL

CONFIDENTIAL

different likelihoods that they arose from the submarine population of contacts. In terms of a dichotomous orientation such as that of the HHIP, the best answer is the same in all three instances: "Submarine." In terms of practicality, since there are distinctly different levels of submarine probability associated with the three cases, the decision maker is denied information that might have an important effect both on his ultimate classification and on his choice of tactical actions.

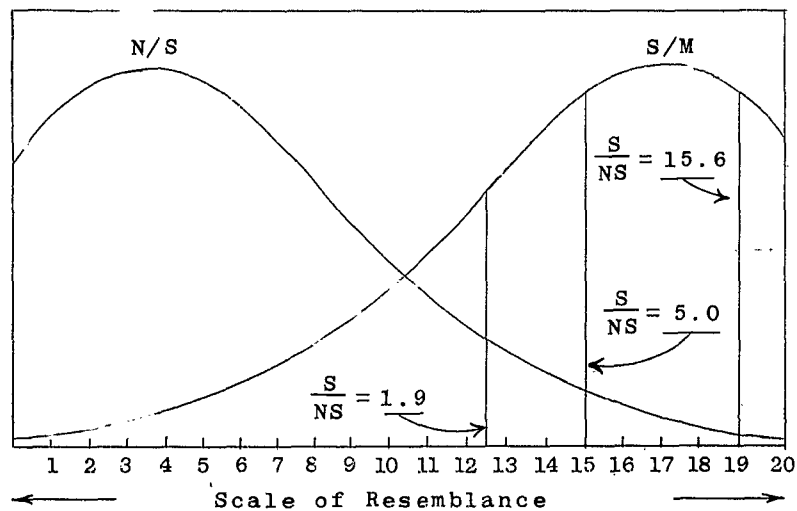


Figure 2. Three observation points on resemblance scale with different frequencies of occurrence for submarine and nonsubmarine clue combinations.

The output of the HHIP is in the form desired by many decision makers. To the extent that other intelligence bearing on the classification decision is lacking, or to the extent that sonar information is weighted more heavily than other sources, the output of the HHIP eliminates the need for a reasoned decision. Naturally the device was not intended to render this kind of service. Like other devices of its kind, it was developed to serve as an aid to

CONFIDENTIAL

CONFIDENTIAL

the classifier--not as a substitute for him. The fact remains, however, that its output is in a form welcomed by some because it relieves them of a burdensome decision-making task.

Whether or not this is a desirable state of affairs probably cannot be decided in the absence of cost considerations such as the mission of the ship, the number and types of weapons available, the consequences of a submarine penetration and so forth. In the terms of the decision theorist, the desirability of a simple dichotomy, or any other format, cannot fully be determined in the absence of a payoff matrix.

The MITEC 20 Final Lights

In the design of MITEC, it was decided that some attempt should be made to display the differential submarine probabilities resulting from the clue correlations although it was recognized that these probabilities could not be established very precisely. As in the case of HHIP, the programming of MITEC also was based on empirically determined likelihood ratios derived from recorded sonar contacts. Additionally it incorporated a logical analysis of the kinds of clue configurations that should be expected from submarines in view of their physical and behavioral characteristics. This latter is necessary in any clue evaluation device at the present time because of the limited samples of recorded sonar contacts available for analysis.

It was decided to attempt to display the point at which each clue configuration fell on the hypothetical scale of submarine resemblance. Although this scale was felt to be continuous, engineering convenience and limited number of qualitatively different solutions discriminable by the MITEC program, suggested that a series of 20 discrete lights would represent the scale adequately. Remarkably little display design guidance was available on the subject and there was little to recommend more elaborate alternatives.

CONFIDENTIAL

CONFIDENTIAL

The 20 lights thus formed an ordinal scale of resemblance, with high numbers assigned to clue configurations having a high probability of having resulted from a submarine and low numbers assigned to configurations having a low probability of having resulted from a submarine. Lights near the middle of the scale were descriptive of targets that had some of the characteristics of both submarines and nonsubmarines (i.e., likelihood ratios near 1.0). The arrangement of lights used in the MK. I Mod. 1 MITEC is shown in Figure 3.

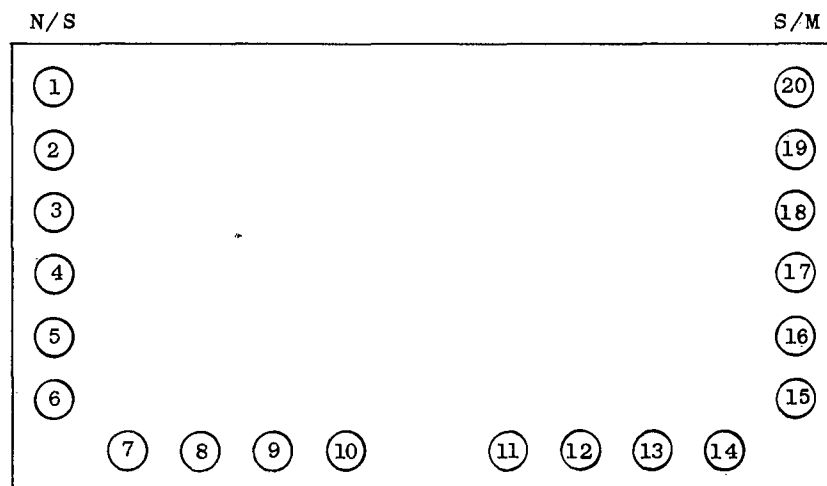


Figure 3. Format of 20-lights display used with MITEC MK. I Mod. 1.

The lights were also color-coded in such a way as to be relatable to the traditional Navy Trichotomy: red lights (15 through 20) were meant to imply that the target was most likely of submarine origin; orange or yellow lights (7 through 14) implied "possible" submarine but with maximum uncertainty; green lights (1 through 6) implied that the target most probably was a nonsubmarine.

It is to be emphasized that the resulting scale was by no means intended as a linear, equal-interval probability scale. Such

CONFIDENTIAL

CONFIDENTIAL

a scale, while highly desirable, of necessity would have had to have been based upon the results of very large samplings of sonar contacts that were representative of the total universe of sonar returns, both submarine and nonsubmarine.

It was known, however, from studies performed during the development of MITEC (5), that a monotonic relationship existed between the value of the MITEC lights and the probability that the corresponding clue configuration resulted from a submarine. It was also known that this relationship was non-linear. A distinct possibility existed that there were better ways to display the information for decision-making purposes. Since, according to detection theory (6), the optimum decision criterion is based on likelihood ratio, it could be hypothesized that a display of the outcome in a form more closely related to likelihood ratio would produce a more optimum set of decisions.

In the absence of large representative samples of target data, objectively establishing the values of the likelihood ratios along the resemblance scale was not feasible. But the meanings conveyed by a likelihood display, in possible contrast to the 20-lights display, could be studied by making some reasonable assumptions about the shapes of the submarine and nonsubmarine frequency distributions along the scale. The interpretation of the displays by the decision maker then would become the criterion for determining whether one display or another was likely to lead to the more optimum decisions. It was with this orientation that the present study was conducted.

THE EXPERIMENTAL QUESTION

The present investigation was designed to answer the following question: How do several alternative techniques for displaying the probable nature of the target to the classifier (decision maker)

CONFIDENTIAL

CONFIDENTIAL

affect his level of certainty as to the nature of the target?
Specifically, what meanings do the following types of information displays have to military personnel who typically make classification decisions?

1. A numerical display of relative frequency of occurrence
2. A bar graph display of relative frequency of occurrence
3. A numerical display of relative odds
4. The traditional Navy classification trichotomy
5. A 20-lights display similar to MITEC

CONFIDENTIAL

CONFIDENTIAL

II. PROCEDURE

Hypothetical Distributions of Submarine and Nonsubmarine Targets

As previously emphasized, there exists no adequate sample of recorded sonar contacts that can be considered truly representative of the universe of sonar targets. It seems reasonable to assume, however, on the basis both of logic and the available evidence (5), that clue configurations from representative samples of submarine and nonsubmarine contacts would distribute themselves on a continuum of resemblance in the overlapping manner depicted in Figures 1 and 2. However, the empirical data suggest that a much greater degree of overlap exists in the center of the scale than that shown in the figure.

It should be recognized that the degree of overlap obtained from any particular sample is a function both of the representativeness of that sample and the discriminability of the sensor/processor system in use. Since the samples of contacts that have been recorded and analyzed are comparatively small, and since they were purposely selected to emphasize the area of overlap (i.e., it was known that they were not representative), it was considered that the distributions assumed in Figure 1 were not unreasonable for the purposes of the present study.

In examining the frequency curve of submarine clue configurations on the resemblance scale, one might be struck by the apparent logical inconsistency of having lower S/M frequency values for very high points on the scale of resemblance (e.g., 19, 20) than for points in the upper mid-range (e.g., 16, 17). This inversion in relationships is sensible, however, when it is remembered that environmental conditions will often, if not usually, operate to degrade the signal characteristics when a submarine is in fact present. Strong but imperfect resemblance is more likely to occur than near-

CONFIDENTIAL

perfect resemblance. This raises questions about the applicability of signal detection theory to the classification problem but the basic concepts still appear useful. It also emphasizes the admitted fact that the resemblance scale is complex and in the real world is discrete rather than continuous.

Whatever the shortcomings of the curves in Figure 1, they did provide a basis for the selection of observation points along the scale of resemblance that would be illustrative of solutions obtainable for a variety of sonar contacts. Further, the output could also be displayed in a form directly related to likelihood ratio.

Selection of Observation Points on Scale of Resemblance

For convenience, a total of 12 observation points was selected from the 20 depicted on the abscissa in Figure 1. Six of these appeared on either side of a 1.0 likelihood ratio.⁶ The 12 points were chosen in such a way that both extremes of the scale were well represented as were all adjacent points having substantially different relative frequencies of occurrence for submarine and nonsubmarine configurations. The values of the ordinates for the submarine and nonsubmarine curves at each selected solution point next were estimated by simple measurement. These, and the corresponding likelihood ratios are given in Table I.

⁶To simplify correspondence with the 20 MITEC lights, the scale of resemblance was divided into 20 equal-sized intervals.

CONFIDENTIAL

CONFIDENTIAL

Table I. Estimated Frequencies (f) at 12 Observation Points and Corresponding Likelihood Ratios (l) for Hypothetical S/M and N/S Contact Distributions

Observation Point	f_{SM}	f_{NS}	$l = \frac{f_{SM}}{f_{NS}}$
1	1.0	10.0	0.10
3	1.2	15.0	0.08
5	1.9	17.2	0.11
6	2.5	17.3	0.14
8	5.4	16.3	0.33
10	11.5	14.0	0.82
12	15.6	10.6	1.47
14	18.3	6.5	2.82
16	20.2	3.0	6.73
17	19.5	1.8	10.82
19	14.0	0.9	15.55
20	9.0	0.8	11.23

Alternative Methods of Display

For each of the selected observation points along the scale of resemblance, five alternative methods of display were designed (see Figures 4a through 4e). These displays have been depicted as they would appear for observation point #8 in Table I.

Figures 4a through 4c were direct reflections of the likelihood ratio information. Figure 4a indicates the Relative Frequency of occurrence for submarine and nonsubmarine clue configurations in simple numerical form.

CONFIDENTIAL

CONFIDENTIAL

Submarine	5
Nonsubmarine	16

Figure 4a. Relative Frequency of occurrence display.

Figure 4b, the Bar Graph, presented the same information contained in the Relative Frequency display but in a familiar graphic form.

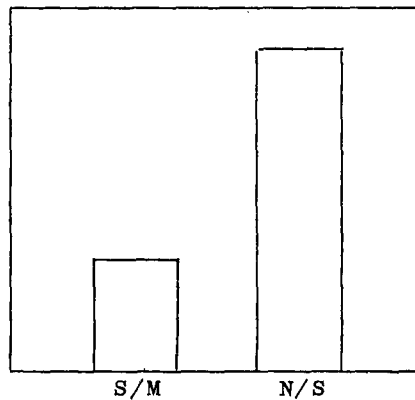


Figure 4b. Bar Graph display.

Figure 4c, Relative Odds, was a third presentation of the basic information available in likelihood ratio. In this case the relative frequency data were simply reduced so that the value 1 appeared either in the numerator or denominator, depending on which class of target was more frequently associated with the observation point in question.

CONFIDENTIAL

Submarine	1
Nonsubmarine	3.2

Figure 4c. Relative Odds display.

In Figure 4d, the solution is shown as it might have been produced by a 20-Lights display such as that of MITEC MK I. Mod. 2. The light values were assigned so that they bore a one-to-one correspondence with the observation points on the scale of resemblance. Actually, the MITEC lights are programmed somewhat differently, a fact more fully discussed later. As indicated, the color coding used with the MITEC as a basis for relating the solution to the traditional Navy trichotomy, also was preserved.

N/S

S/M

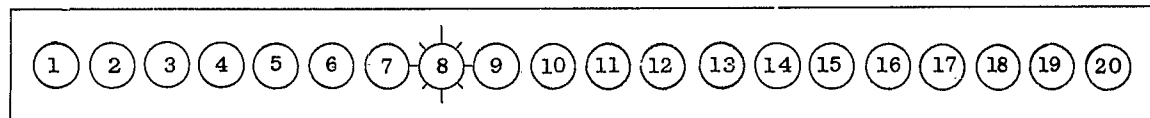


Figure 4d. 20-Lights display.

Finally, a display was presented as it might occur in classifying according to the Navy Trichotomy (see Figure 4e). The decision as to which of the three categorical solutions to display for each observation point was made on the basis of the 3-way division also made for MITEC: points 1 through 6 were displayed as "Nonsubmarine"; 7 through 14 were displayed as "Possible Submarine"; and 15 through 20 were displayed as "Probable Submarine." While there was no basis for expecting an interpretive correspondence, this technique assured that method of display would be encountered equally often by the subjects and provided a means of determining the meaning of the

CONFIDENTIAL

CONFIDENTIAL

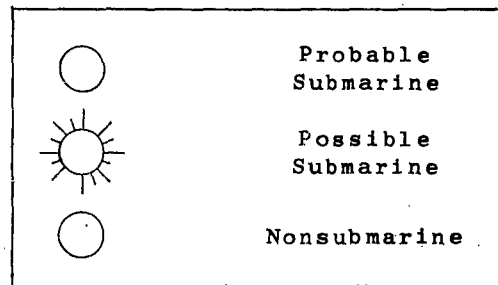


Figure 4e. Navy Trichotomy display.

Trichotomy reliably compared to the other displays.

This procedure was followed for all 12 stimulus conditions (observation points) and each of the 5 alternative displays making a total of 60 display conditions. Each of the 60 display conditions was produced on a single sheet of paper, arranged into random order, and assembled into a test booklet.

The Response Scale

Ideally, it would have been desirable to learn whether the several alternative methods of display differentially affected the decision maker's choice of action. This would require a pay-off matrix based on a number of hypothetical tactical situations representing several levels of threat. In discussions with Naval officers, it became evident that it was a practical impossibility to specify many tactical situations in sufficient detail to evoke reasonably uniform evaluations of the threat imposed.

Since this was not basically a study of the decision-making process, but rather an investigation of the meaning conveyed by several alternative displays of the same classification observation, it was decided to avoid this complication by eliciting a relatively simple statement of meaning rather than one concerning the more involved criterion of action.

CONFIDENTIAL

Because all sonar target classifications are made with something less than complete certainty, it seemed reasonable to ask the subjects for a statement of their level of certainty or uncertainty for each selected display condition. This was accomplished by use of the simple graphic rating scale depicted in Figure 5. For each display-observation condition, the subjects were asked to indicate how

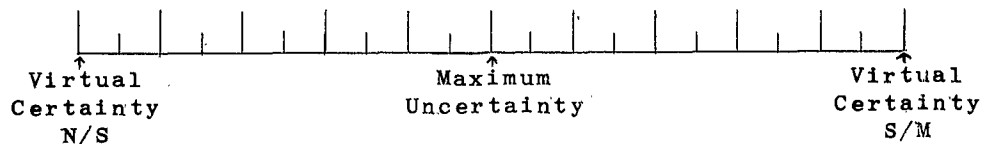


Figure 5. Response scale for indicating level of certainty.

certain they felt that the target was of submarine or nonsubmarine origin by placing a check mark anywhere along the line that appropriately described their level of certainty. The verbatim instructions given to the subjects were as follows:

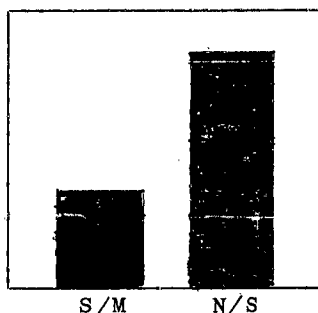
"It is well established that the nature of a sonar contact (whether it is a submarine or nonsubmarine) cannot be determined with certainty solely on the basis of the target information generated by sonar. This fact is recognized in the traditional manner in which sonar reports the contact to command as 'probable submarine,' 'possible submarine,' or 'nonsubmarine.' All classification solutions from sonar are probabilistic, that is, they involve some degree of uncertainty.

"Now it is possible to report sonar's classification to command in a variety of ways in addition to the traditional one. Whether or not some of these ways are desirable naturally depends on whether they lead to appropriate and consistent interpretations by command. The purpose of this meeting is to test the meaningfulness of several of these alternatives.

CONFIDENTIAL

CONFIDENTIAL

"On the pages of the booklets that will be passed out to you, will be found the results of the analysis of a number of sonar contacts. These results are displayed in several ways for each contact. For example, you may find a display that looks like this:" (Figure below was drawn on blackboard.)



"Underneath the display will be found a 'scale of certainty' which looks like this." (Figure 5 was drawn on board.) "We would like you to interpret the results in each display by placing a check (✓) anywhere along this scale of certainty that appropriately expresses your assessment of the contact. For example, in this case you might place the check here." (A check-mark was placed approximately half-way between the left end and the mid-point of the response scale.)

"Are there any questions?"

"Please work as rapidly as you please. We are most interested in your immediate impression rather than in a calculated response."

The Subjects

Eleven Naval officers ranging in rank from Lieutenant (j.g.) through Captain, and 12 petty officers, First Class and Chief, served as subjects. All were on the instructional staff or in tactical courses at the Fleet ASW School, San Diego, at the time of the study.

CONFIDENTIAL

CONFIDENTIAL

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

Their responses on the scale of certainty were quantified by a linear translation of the distance of the check-mark from the left-hand end of the scale (Virtual Certainty N/S). Since the numerical values were arbitrary, the left end of the scale was taken as zero, the right end as 20.0. This had the advantage of making the values directly relatable to the 20-Lights display. The subjects, of course, had no knowledge of the score values to be used although they could have ascertained, if they had taken the trouble, that the response scale had 20 index points.

CONFIDENTIAL

CONFIDENTIAL

III. RESULTS

Expressions of Certainty

Inspection of Figure 6 indicates that the three alternative methods of display based on likelihood information (Relative Frequencies, Relative Odds, and Bar Graph) produced essentially the same judgmental function. Moreover, all three functions fairly closely approximated that which would be generated by the theoretical ideal observer, i.e., an individual whose level of certainty (and presumably, decisions) corresponded directly with the relative magnitude of the likelihood ratios. The behavior of the ideal observer is represented in Figure 6 by the solid line.

There are some interesting departures from the ideal function, however, particularly for the more extreme observation points. In general, the level of certainty is less than that justified by the likelihood ratios for points 1 through 6 and 14 through 20. This conservative response may reflect a logical consideration on the part of the subjects that "virtual certainty" is never achievable on the basis of sonar data alone. Data from the 20-Lights display, which do not show this effect, raise doubts as to the validity of this explanation, however. Edwards (8, p. 16) presents evidence that subjects are often unwilling to extract from information all the certainty that is in it. Apparently the present finding supports his observations.

Inspection of Table II reveals some additional interesting comparisons. For example, although the likelihood ratio for observation point 19 was substantially different from those for points 17 and 20, no uniform reflection of this difference is seen in the certainty ratings. The Relative Odds display produced notably uniform results for all three points. Similar departures from the ideal can be seen for observation points 6 and 8. Although the

CONFIDENTIAL

CONFIDENTIAL

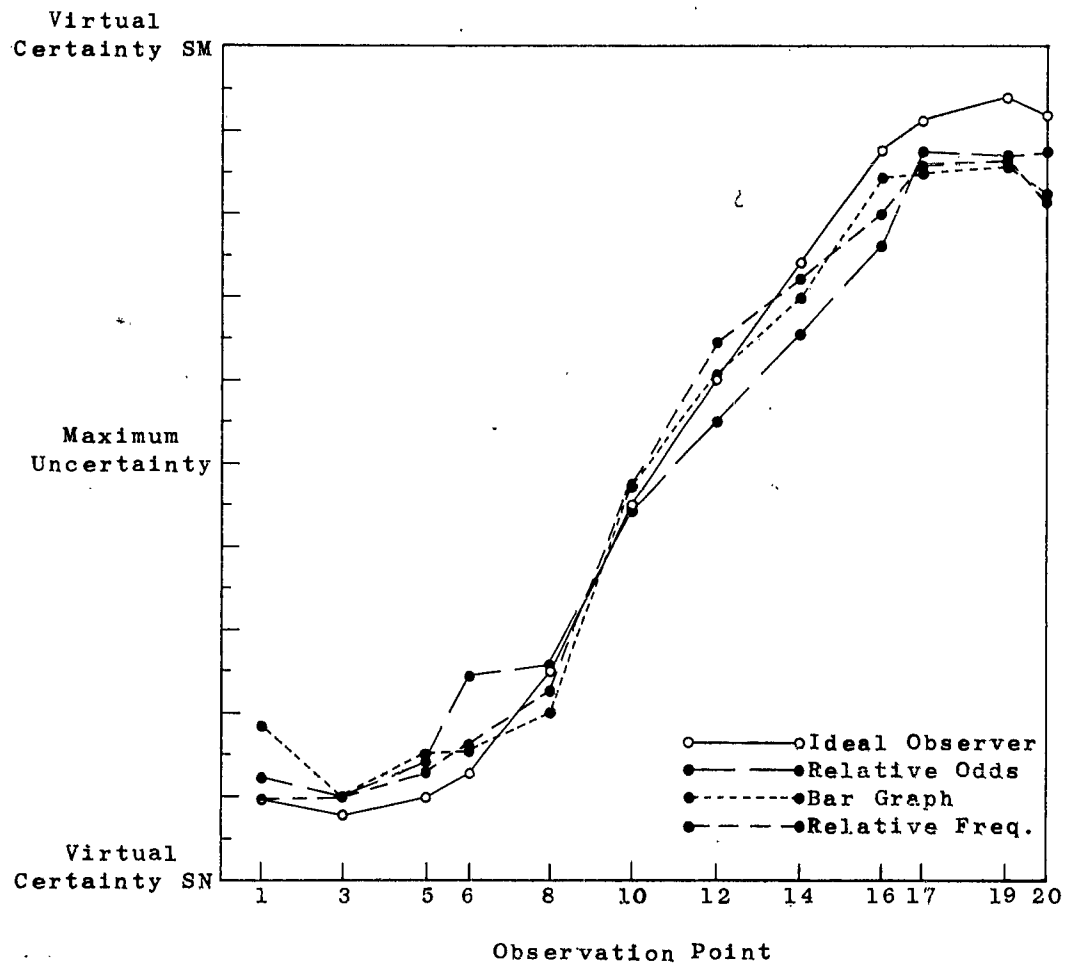


Figure 6. Certainty ratings produced by three displays of relative likelihood. (Each point based on 23 ratings.)

CONFIDENTIAL

CONFIDENTIAL

Table II. Average Certainty Ratings for Five Displays
Related to Theoretically "Ideal" Observer

(N = 23)

Observation Point	1	3	5	6	8	10	12	14	16	17	19	20
Likelihood Ratio	0.10	0.08	0.11	0.14	0.30	0.82	1.47	2.82	6.73	10.82	15.55	11.23
"Ideal" Observer	9.0	7.5	10.0	12.5	25.0	45.0	59.5	74.0	86.5	91.5	94.0	92.0
1. Relative Frequencies	9.8	9.9	13.1	15.9	22.1	47.6	64.0	72.0	79.4	86.1	86.6	81.9
2. Bar Graph	17.4	10.0	14.9	15.5	19.6	46.8	60.1	70.3	84.9	85.2	86.7	82.5
3. Relative Odds	12.0	10.1	13.0	23.9	25.4	44.4	54.5	65.9	76.0	87.5	87.3	87.4
4. 20-Lights	4.2	14.4	20.2	26.3	38.0	50.3	59.0	68.2	76.2	78.7	90.5	94.8
5. Navy Trichotomy	13.1	14.5	14.7	15.0	58.5	56.3	54.6	57.3	75.8	80.7	83.0	80.1
"Nonsubmarine"				"Possible Submarine"				"Probable Submarine"				

CONFIDENTIAL

CONFIDENTIAL

response curves generally approximated that of the ideal observer, there were frequent departures at individual data points. There appeared to be no systematic pattern to these departures. Possibly sequential and contrast effects were operative.

Comparison of the three displays with each other is suggestive of other possible differences in the meanings conveyed. A number of statistically significant differences are apparent for particular data points. For example, at observation point 1, the Bar Graph produced significantly less certainty ($p < .01$) than the Relative Frequency display. At observation points 6 and 12, the Relative Odds display produced significantly less certainty than the Relative Frequency display ($p < .01$) and it appears generally to have produced less certainty than the other displays throughout the middle observation points.

Figure 7 depicts the certainty levels elicited by the 20-Lights display in comparison to the ideal observer. Not surprisingly, perhaps, this display was interpreted by the observers as an essentially linear, equal interval, probability scale. With the exception of the two extreme observation points (p_1 and p_{20}) and the approximate mid-point (p_{12}), this display elicited uniformly conservative responses. Inspection of the dispersion of certainty ratings in Figure 7 reveals that overly conservative responses were obtained in general for points 3 through 10 and 14 through 19. In comparison to the three likelihood displays, the 20-Lights display also produced significantly more conservative responses, again with the notable exceptions at the two extreme observation points.

In view of the actual programming of MITEC, which employs a 20-lights final display, it may be concluded from these results that current models of MITEC are also likely to produce overly conservative evaluations. While the shapes of the underlying distributions of submarine and nonsubmarine target clue combinations

CONFIDENTIAL

CONFIDENTIAL

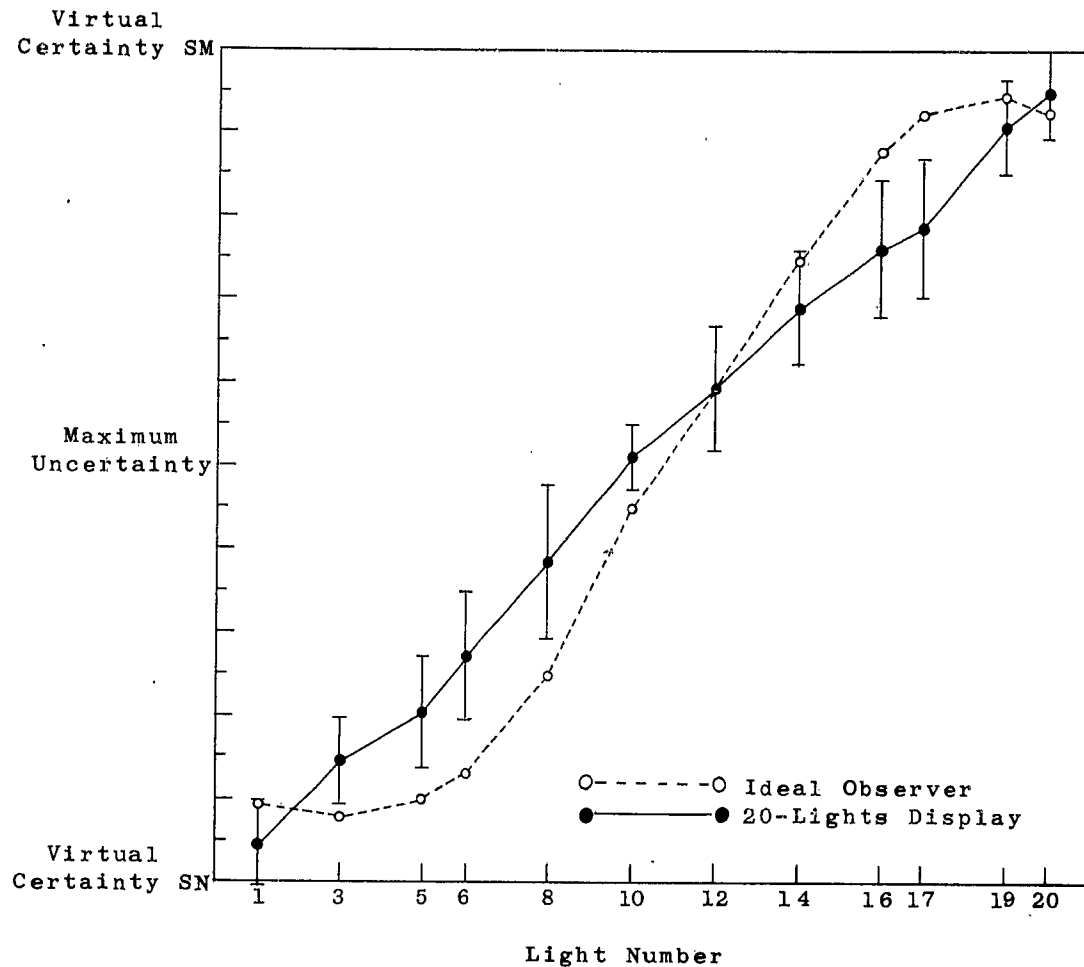


Figure 7. Mean certainty and dispersion of ratings produced by the 20-Lights display (23 observers).

in the real world are not fully known because of incomplete samples, it is likely that they are somewhat similar to those assumed for this study. There is some empirical evidence, moreover, that extremely low probabilities of submarine obtain for clue combinations resulting in Light 5 or lower, and extremely high probabilities obtain for combinations resulting in Light 16 or higher (5).

CONFIDENTIAL

CONFIDENTIAL

Figure 8 depicts the certainty function produced by the traditional three-category Navy classification of sonar targets in relation to the ideal observer. This display was based on a categorical, rather than a continuous function, and, with only three data points, could not be related directly to the likelihood ratios used for the other displays. However, it appeared meaningful to use the color coding of the 20-Lights display as the basis for

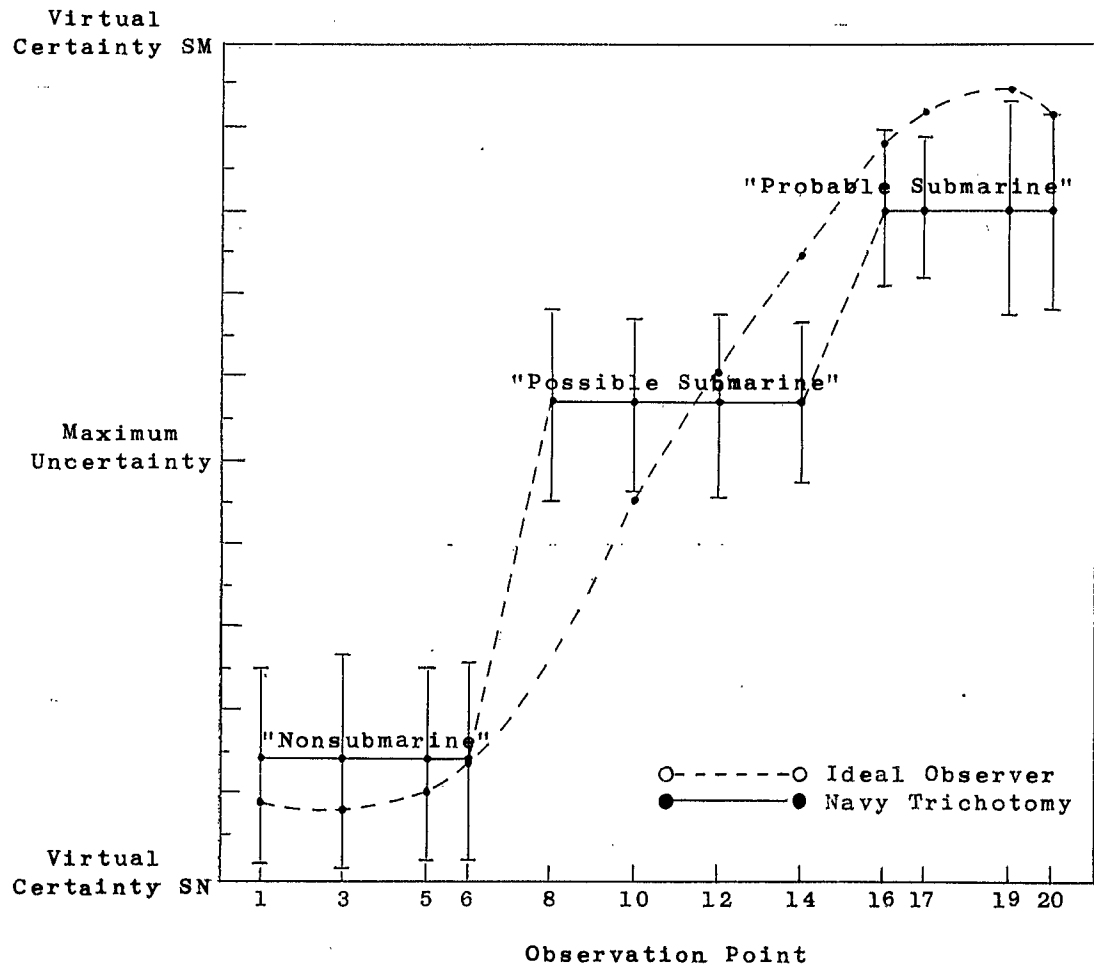


Figure 8. Mean certainty and dispersion of ratings produced by traditional Trichotomy display (23 observers).

CONFIDENTIAL

relating the output of the traditional classification system to the observation points used for this study. It will be recalled that the 20-Light display preserved the suggestion of a trichotomous output by employing green, amber, and red lights for clue combinations showing little, some and much resemblance to the submarine. When this was done, the step-like function of Figure 8 was generated. The numerical data that generated this function are reported in Table II.

It will also be recalled that for the sake of balance each stimulus condition for the Trichotomous display was presented four times. The results are given as four separate estimates of the mean certainty level for each categorical classification. It will be noted that the group response to each presentation was essentially the same.

The average certainty rating for the statement "probable submarine" was 79.9. This is most comparable to the certainty ratings produced by observation point 16 for the likelihood displays, corresponding to a likelihood ratio of about 6.7. It compares with observation point 17 for the 20-Lights display. This seems a reasonable result when it is remembered that the Trichotomous display, according to instructions, was to be considered as reflecting the output of sonar alone. On the other hand, the official criteria for the statement "probable submarine" are quite stringent (see p. 5). If all were met, it could be argued that it would be reasonable to expect a much higher level of certainty than that obtained.

In contrast to the ideal observer, the level of certainty elicited by "probable submarine" is highly conservative. It is evident that much useful discriminative information would be lost to the decision maker if it were necessary to collapse a number of observation points (16 through 20 in this example) into one such

CONFIDENTIAL

CONFIDENTIAL

qualitative conclusion. Further, for the target distributions assumed, systematically overly conservative conclusions would result.

The average level of certainty for the statement "possible submarine" (56.7) was about midway between that elicited by observation points 10 and 12 on the other displays. This is quite close to a likelihood ratio of 1.0 and consequently to a maximally uncertain response. This result suggests that a classification in this category provides minimum information useful to the decision maker and, in practice, probably serves mainly to stimulate continued investigation. Again, the lack of discrimination among quite different configurations of target clues that would occur with this overly simplified categorical conclusion is evident.

The average level of uncertainty for the statement "nonsubmarine" (14.3) was most comparable to the response elicited by observation point 5 for the likelihood displays and points 3 and 5 for the 20-Lights display. The corresponding likelihood ratio would be 0.11 or less. This classification thus appears more decisive than "probable submarine" although it still reflects less certainty than the ideal observer for observation points 1 through 5.

Differences in Interpretation Between Subjects

The dispersion of certainty ratings about the average value at each observation point reflects the level of agreement or similarity of interpretation for each display. Obviously an effective display will minimize individual differences in interpretation and the ideal observer will be perfectly reliable. The data are reported in Table III.

CONFIDENTIAL

CONFIDENTIAL

Table III. Inter-Judge Agreement as Reflected by
Dispersion of Certainty Ratings (σ)

(N = 23)

Observation Point	1	3	5	6	8	10	12	14	16	17	19	20	Aver- age σ
Likelihood Ratio	0.10	0.08	0.11	0.14	0.30	0.82	1.47	2.82	6.73	10.82	15.55	11.23	
1. Relative Frequencies	7.1	6.7	8.9	8.6	9.0	6.9	11.0	10.0	9.0	7.5	8.1	9.2	8.5
2. Bar Graph	5.9	4.4	4.9	6.1	10.9	5.5	7.1	8.3	6.6	8.3	7.0	9.8	7.1
3. Relative Odds	7.1	7.2	7.1	8.8	10.6	10.5	5.5	14.4	9.6	6.8	9.2	8.0	8.7
4. 20-Lights	4.9	5.2	6.4	7.8	8.9	3.9	7.0	6.3	8.2	8.1	5.8	5.2	6.5
5. Navy Trichotomy	12.0	12.1	11.0	11.2	11.7	10.6	11.0	9.8	9.8	8.6	12.5	12.0	11.2
Average σ	7.4	7.1	7.7	8.5	10.2	7.5	8.3	9.8	8.6	7.9	8.5	8.8	

CONFIDENTIAL

CONFIDENTIAL

It is evident in general that the 20-Lights display produced the greatest agreement. This was true particularly at the two extreme observation points (1 and 20) and near the middle (point 10). These points probably were more effectively portrayed on the 20-Lights display than on the others.

The average σ of the 20-Lights display was significantly lower ($p < .01$)* than all other displays except the Bar Graph. The Bar Graph in turn produced greater agreement ($p < .05$) than the Relative Frequencies and Relative Odds displays. Finally the Traditional Trichotomy produced significantly more disagreement than all of the other displays ($p < .01$). It seems likely that this was a function of the relative ambiguity of verbal as opposed to numerical or graphic symbols.

For the most part the magnitude of the dispersions did not show any systematic trends across the several observation points. The mean dispersion was greatest at points 8 and 14 which, curiously, involve similar likelihood ratios of about 1:3 (or 3:1). It may be that moderate odds, as opposed to extreme or nearly equal odds, produce the greatest interpretive differences although the trends are not sufficiently consistent in the present study to be more than suggestive.

A few of the unexpectedly large σ 's were probably due to effects of position in the stimulus series. For example, Bar Graph point 8 ($\sigma = 10.9$) was the very first stimulus in the booklet. Relative Odds point 14 ($\sigma = 14.4$) possibly involved a contrast effect, the immediate preceding stimulus by chance being Relative Odds point 16.

Differences in Interpretation Between Officers and Petty Officers

There were small but consistent differences between the Officers' and Petty Officers' expressions of certainty for the

*According to the sign-rank tests of Wilcoxon as described by Guilford (9).

CONFIDENTIAL

several displays. These are summarized in Table IV. The differential confidence is shown in relation to the point of maximum uncertainty since the Officers were consistently more certain of a nonsubmarine likelihood and the Petty Officers were consistently more certain of a submarine likelihood. As indicated, these biases were not symmetrical about the point of maximum uncertainty and varied in magnitude for the several displays.

Table IV. Differential Confidence as to Target's Nature

Display	Officers	Petty Officers
1. Relative Frequencies	3.7% NS	2.8% SM
2. Bar Graph	1.8% NS	2.7% SM
3. Relative Odds	2.0% NS	0.6% SM
4. 20-Lights	1.0% NS	4.2% SM
5. Navy Trichotomy	1.4% NS	2.5% Prob. SM

It seems likely that these differences are attributable to a somewhat greater skepticism among officers as to the ability of sonar to resolve the nature of a target. Perhaps they were more mindful of the several criteria that must be met before a target can be considered likely to be of submarine origin according to doctrine.

CONFIDENTIAL

CONFIDENTIAL

IV. DISCUSSION

Display Design and the Evaluation Function

It is surprisingly difficult to obtain agreement from Naval Officers concerning the quantity of information desired, or the form in which it is desired, for making a target classification decision. There appear to be two schools of thought. One, whose members apparently refuse to admit the probabilistic nature of sonar evidence, desires a simple dichotomous statement that the target is, or is not, a submarine.* The second school recognizes that even if the statement is made in dichotomous (or trichotomous) form, there are bound to be different levels of certainty associated with the variety of target contacts so classified. Therefore, it feels that some statement of certainty should be required and in addition, may feel a need for supportive evidence in more or less detail, possibly to the point of reporting the specific nature of clues that are displayed.

Essentially, the output of the HHIP is of the dichotomous form desired by the 'black and white' school. Undoubtedly, its simple format will result in considerable acceptance for the reasons discussed above.** There seems little doubt, however, that considerable information of value to a tactical commander is lost if no basis is provided for discriminating between a target that displays most of the properties of a submarine and one that displays just enough of them so that it cannot be dismissed.

*Though these persons are "anti" probability statements, it might be conjectured that if they were contemplating a picnic, their decision to go or not might be based upon whether the forecast was "Probability of rain, 60%," or "Probability of rain, 90%."

**Only the display characteristics of the HHIP are under discussion here. Its considerable usefulness as a classification aid has been demonstrated.

CONFIDENTIAL

In spite of this, it must be argued that a dichotomous submarine-nonsubmarine display would be the ideal one if the sensors discriminated sufficiently, and reliably, between the two classes of targets. In such a case, with essentially zero overlap between target classes on the scale of resemblance, the sonar evidence would be conclusive and the tactical decision would be more or less determined also. The difficulty, of course, is that present sensors do not operate in this discriminating fashion.

In addition to sonar's evaluation, the 'skeptics' school wants as much detail as reasonably can be conveyed about the nature of a target. They should also be given an objective means of evaluating the detailed evidence. With the present limited capability of sonar sensors and signal processing techniques, the conclusion seems inescapable that this group, given some display of the likelihood function, will produce the greater number of effective decisions.

In the MITEC, in addition to the 20 "final" lights, an attempt was made to provide a kind of supportive evidence by employing an "Intermediate Display" that summarized the major characteristics of the target as determined by clue correlation. The sensed size, aspect, reflectivity, depth and motion of the target were reported. Since these are the characteristics that determine the target's degree of resemblance to a submarine, the decision maker was provided with a basis for analyzing in some detail why a particular contact did or did not produce a relatively certain solution.

The greater the degree of overlap between the submarine and nonsubmarine distribution functions on the scale of resemblance, the more important it would seem to provide the decision maker not only with a statement of the relative probabilities involved but with supportive evidence indicating how those values were determined. The less certain the several sources of critical evidence, the more necessary it would seem that the decision maker be appraised of the kind and quality of information obtained.

CONFIDENTIAL

CONFIDENTIAL

Display Design and Tactical Flexibility

The previous line of reasoning can be extended to the decision-making problem faced by officers in tactical command of a number of different types of ASW units. These officers should be apprised of the relative quality and reliability of the information produced by different classes of vehicles (aircraft, surface ships, submarines) with their different sensors (radar, MAD, sonobuoys, active sonars, passive sonars, etc.). Further, officers responsible for the tactical control of vehicles such as the P3V, with its several sensor systems (visual, radar, JULIE, JEZEBEL, ECM, etc.), each having different capabilities and limitations as detection and classification systems, must be apprised of these differential capabilities if they are to make the greatest possible number of correct decisions.

It is likely that many tactical decisions involving these systems are made on a very subjective basis at the present time. Unfortunately, few objective performance figures are available for these systems and comparative data for different systems operating against a target under similar conditions are practically unknown. While differences in detection capability are somewhat predictable, one can only guess at the differential classification capabilities involved.

If it is agreed that the present state of the art does not warrant a dichotomous classification display, the argument for a probabilistic display can be advanced further on the basis of the tactical flexibility it would afford. The question of whether 5, 10, 20, or more levels of likelihood should be displayed is moot. It would seem reasonable, however, that the classification display should provide as many levels of solution as the clue correlator can reliably discriminate. If a given configuration of clues is reliably associated more frequently with one class of target than

CONFIDENTIAL

CONFIDENTIAL

another, then the implied differential probability should be conveyed to the decision maker. The only apparent difficulty is that no sufficiently representative sample of target contacts exists that would enable the differential probabilities to be established empirically.

This problem was circumvented in the MITEC, apparently with reasonable success, by supplementing the empirical target data available with logical expectation. Since the limits of size, motion, aspect, depth and reflective pattern can be reasonably specified for submarines, it was possible to adjust the solution appropriately for various combinations of these attributes even though many combinations actually had not been encountered. The reasonableness of this process was subsequently demonstrated when it was shown that submarine and nonsubmarine contacts did distribute themselves toward opposite ends of the resemblance continuum in a fashion similar to that expected from two overlapping but partially differentiable target classes (5). Thus the several levels of solution displayed by the MITEC 20-lights do provide a basis for establishing differential probabilities that the target is a submarine but there is no assurance (in fact it is highly unlikely) that these probabilities increase in the linear, equal-interval fashion that the subjects of this study expected.

Designing the Displays vs. Programming the Output

While it appears that decision makers will increasingly be confronted with systems that produce probabilistic information, no clear-cut design guidance for the display of such information, either from the human engineering literature or elsewhere, has proved to be readily available. The attempts of the present study were admittedly limited and perhaps lacking in imaginativeness.

The problem is more readily solved, of course, if the underlying

CONFIDENTIAL

CONFIDENTIAL

probability density distributions and the amount of their overlap in the discrimination space is known. As one display possibility, the actual overlapping distributions could be portrayed and the appropriate observation point for any given clue configuration could be indicated directly. This would be a most informative display although its proper interpretation might require considerable training. However, it seems unlikely that target information will be available in the form required by this approach for most military systems. It is not now available for ASW systems and present data collection programs are far too limited to give hope of determining the precise form of these distributions soon. Further, any marked change in system characteristics will result in changes in either the shape, or degree of overlap, of these distributions, or both.

In view of these problems, it would seem that the best design guidance must come from investigations of the meaning of probabilistic displays of various designs to the decision makers. Ideally such studies should provide the opportunity actually to choose between alternative tactical actions. This implies a specifiable pay-off matrix which is not realistically attainable for many, if not most, military problems.

In the absence of action criteria, a useful compromise would seem to be the employment of intermediate criteria such as those used in this study--measures of certainty and agreement. The design problem then becomes one of achieving an interpretive match between the meaning of the display and the discriminating capabilities of the sensing and processing equipment. For example, had it been known that the 20-lights display of MITEC produces an essentially linear, equal-interval scale of certainty, a closer match could have been achieved between the programmer's intent and the observer's interpretation of the relative likelihoods involved. Whatever the system of probabilistic information processing, considerably more knowledge of the interpretations placed by typical

CONFIDENTIAL

CONFIDENTIAL

observers on the display of such information is desirable if the decision process is to be aided effectively.

CONFIDENTIAL

CONFIDENTIAL

REFERENCES

1. Mackie, R.R. Learning problems in sonarmen training, Technical Report No. 3. Classification proficiency using experimentally developed classification principles. (For Personnel and Training Branch, Psychological Sciences Division, Office of Naval Research.) 1955, Contract Nonr 1106(00) CONFIDENTIAL
2. Mackie, R.R., Gavin, R.A. and Parker, E.L. Target classification using active scanning sonars. (For Personnel and Training Branch, Psychological Sciences Division, Office of Naval Research.) 1959, Contract Nonr 2649(00) CONFIDENTIAL
3. Gavin, R.A. Manual for use with hand-held information processor. (For Bureau of Ships, Department of the Navy.) September 1961, Contract NObsr 85425 CONFIDENTIAL
4. Mackie, R.R. and Harabedian, A. Learning problems in sonarmen training, Technical Report No. 10. A study of three methods of processing sonar information for target classification. (For Personnel and Training Branch, Psychological Sciences Division, Office of Naval Research.) 1958, Contract Nonr 1106(00) CONFIDENTIAL
5. Mackie, R.R. and Gavin, R.A. The Mitec Mk. 1 Mod. 0 target classification aid (U). (Development, performance, and technical manual) (For Undersea Programs, Code 466, Office of Naval Research, Department of the Navy.) 1962, Contract Nonr 3530(00) CONFIDENTIAL
6. Green, David M. "Psychoacoustics and detection theory." J. acoust. society of America, 1960, 32: No. 10, 1189-1203.
7. Edwards, Ward D. "Dynamic decision theory and probabilistic information processing." Human Factors, 1962, 4, 59-73.
8. Edwards, Ward D. "Probabilistic information processing in command and control systems." Engineering Psychology Laboratory, Institute of Science and Technology, The University of Michigan, March 1963.
9. Guilford, J.P. Fundamental Statistics in Psychology and Education. (3rd Edition) New York: McGraw-Hill, 1956.

CONFIDENTIAL

CONFIDENTIAL

DISTRIBUTION LIST FOR TECHNICAL REPORT NO. 206-21

Note: Send 15 copies to
addressee shown; retain
balance for future requests

35	Chief of Naval Research Code 458 Department of the Navy Washington 25, D. C.	1	Commander Cruiser Destroyer Force U.S. Pacific Fleet San Diego 32, California
10	Chief of Naval Operations Department of the Navy Washington 25, D. C. (1 ea. to:) Op 32 71 Op 10 712 31 713 316 03EG 34 03T 07T	1	Commander Antisubmarine Force U.S. Atlantic Fleet U.S. Naval Base Norfolk 11, Virginia
		1	Commander Operational, Test & Evaluation Force U.S. Atlantic Fleet U.S. Naval Base Norfolk 11, Virginia
1	Commander in Chief U.S. Atlantic Fleet U.S. Naval Base Norfolk 11, Virginia	1	Commander Key West Test & Evaluation Detachment Key West, Florida
1	Commander in Chief U.S. Pacific Fleet Fleet Post Office San Francisco, California	1	Chief of Naval Personnel Pers-C1413 Department of the Navy Washington 25, D. C.
1	Commander Training Command U.S. Atlantic Fleet U.S. Naval Base Norfolk 11, Virginia	1	Chief, Bureau of Ships Code 688 Department of the Navy Washington 25, D. C.
1	Commander Training Command U.S. Pacific Fleet c/o U.S. Fleet Sonar School San Diego 47, California	1	Chief, Bureau of Ships Code 689A Department of the Navy Washington 25, D. C.
1	Commander Destroyer Force U.S. Atlantic Fleet Newport, Rhode Island	1	Commander Antisubmarine Force U.S. Pacific Fleet Navy #128 Fleet Post Office San Francisco, California

CONFIDENTIAL

CONFIDENTIAL

1	Chief, Bureau of Ships Code 370 Department of the Navy Washington 25, D. C.	1	Commanding Officer & Director U.S. Navy Electronics Laboratory Code 3180 San Diego 52, California
1	Chief, Bureau of Ships Code 374 Department of the Navy Washington 25, D. C.	1	Commander Naval Ordnance Test Station Pasadena, California
1	Chief, Bureau of Naval Weapons Code RU221 Department of the Navy Washington 25, D. C.	3	Chief of Naval Research Code 466 Department of the Navy Washington 25, D. C.
1	Chief, Bureau of Naval Weapons Code RU DC11 Department of the Navy Washington 25, D. C.	1	Commander New London Development Detachment Operational Test & Evaluation Force U.S. Navy Underwater Sound Laboratory Fort Trumbull New London, Connecticut
1	Chief, Bureau of Naval Weapons Code RU Department of the Navy Washington 25, D. C.	1	Director Naval Research Laboratory Code 2027 Department of the Navy Washington 25, D. C.
1	Director Naval Research Laboratory Code 1021 Department of the Navy Washington 25, D. C.	1	Commander Destroyer Squadron 12 (DesDiv 121) Fleet Post Office New York, New York
1	Director Naval Research Laboratory Code 5550 Department of the Navy Washington 25, D. C.	1	Chief of Naval Research Code 455 Department of the Navy Washington 25, D. C.
1	Commanding Officer & Director U.S. Naval Training Device Center Port Washington, New York	1	Chief of Naval Research Code 468 Department of the Navy Washington 25, D. C.
1	Commanding Officer & Director Code 900 U.S. Navy Underwater Sound Laboratory Fort Trumbull New London, Connecticut	1	Chief of Naval Research Code 407S Department of the Navy Washington 25, D. C.

CONFIDENTIAL

CONFIDENTIAL

2	Commanding Officer U.S. Fleet ASW School San Diego 47, California	1	Commander Destroyer Flotilla 3 c/o U.S. Naval Station Long Beach 2, California
2	Commanding Officer U.S. Fleet Sonar School U.S. Naval Base Key West, Florida	1	Commander Destroyer Flotilla 4 Fleet Post Office New York, New York
1	Commanding Officer U.S. Fleet Training Center U.S. Naval Base Norfolk 11, Virginia	1	Commander Destroyer Flotilla 5 Fleet Post Office San Francisco, California
1	Commanding Officer U.S. Fleet Training Center U.S. Naval Base Newport, Rhode Island	1	Commander Destroyer Flotilla 6 Fleet Post Office New York, New York
2	Commanding Officer & Director U.S. Atlantic Fleet ASW Tactical School U.S. Naval Base Norfolk 11, Virginia	1	Commander Destroyer Squadron 28 (DesDiv 281) Fleet Post Office New York, New York
1	Commander Fleet Training Group Navy No. 115 Fleet Post Office New York, New York	1	Commander Destroyer Squadron 36 (DesDiv 361) Fleet Post Office New York, New York
1	Commander Fleet Training Group San Diego, California	1	Commander Destroyer Squadron 25 (DesDiv 251) Fleet Post Office San Francisco, California
1	Commander Destroyer Development Group 2 Fleet Post Office New York, New York	3	Chief of Naval Personnel Pers-152 Department of the Navy Washington 25, D. C.
1	Commander Destroyer Flotilla 1 Fleet Post Office San Francisco, California	1	Chief of Naval Personnel Pers-C123 Department of the Navy Washington 25, D. C.
1	Commander Destroyer Flotilla 2 Fleet Post Office New York, New York	1	Commanding Officer ONR Branch Office 1030 E. Green Street Pasadena 1, California

CONFIDENTIAL

CONFIDENTIAL

- 1 Commanding Officer
Office of Naval Research
Navy #100
Fleet Post Office
New York New York
- 1 Director
Naval Research Laboratory
Attn: Technical Information
Officer
Washington 25, D. C.
- 1 Commanding Officer
U.S. Naval Air Development
Center
Johnsville, Pennsylvania
- 1 Superintendent
U.S. Naval Postgraduate School
Monterey, California
Attn: Librarian
- 1 Commanding Officer
Submarine Development Group 2
Fleet Post Office
New York, New York
- 1 Executive Secretary
Committee on Undersea Warfare
National Research Council
2101 Constitution Avenue
Washington 25, D. C.
- 1 Systems Analyses Group
Undersea Warfare Research &
Development Council
Room 5-224
Naval Ordnance Laboratory
Silver Spring 19, Maryland
- 1 Hydrographer
U.S. Navy Hydrographic Office
Washington 25, D.C.
- 1* Mr. Robert Rubega
Underwater Acoustics Section
Research & Advanced Development
Dept.
Stromberg-Carlson Company
100 Carlson Road
Rochester 3, New York
- * Note: Reports through
Confidential be sent to
addressee with copy of the
letter of transmittal to
INSMAT, 740 Main Street,
Buffalo, New York or send
report via INSMAT
- 1 Director, Sonar Research
Division
U.S. Naval Personnel Research
Field Activity
Naval Medical Research
Laboratory
U.S. Naval Submarine Base
New London, Connecticut
- 1 Commanding Officer
U.S. Naval Air Development
Center
Johnsville, Pennsylvania
Attn: NADC Library
- 1 Chief of Naval Personnel
(Pers-C1112a)
Department of the Navy
Washington 25, D. C.
- 1 Commander, Ocean Systems
U.S. Atlantic Fleet
Norfolk 11, Virginia
- 1 Supreme Allied Commander
U.S. Atlantic Fleet
ASW Research Center
APO 19
New York, New York
via 1) ONR Code 458
2) CNO OP 92C2
3) Secretariat of
Military Information
Control Committee
- 1 CDR Jack Adams
Qtrs. P
Navy #128
Fleet Post Office
San Francisco, California
- 1 Security Officer
Hudson Laboratories
145 Palisades Street
Dobbs Ferry, New York
Attn: Dr. Dana P. Mitchell

CONFIDENTIAL

CONFIDENTIAL

SUPPLEMENTARY LIST FOR TECHNICAL REPORT NO. 206-21

2	Office of Naval Research	Commanding Officer and
1	Department of the Navy	Director
	Washington 25, D. C.	Naval Research Laboratory
	Attn: Code 406T	Washington 25, D. C.
	Code 467	Attn: Code 2000
		Code 5550
10	Director	
	Armed Forces Technical	1 U.S. Navy Underwater
	Information Agency	Ordnance Station
	Arlington Hall Station	Newport, Rhode Island
	Arlington 12, Virginia	
	Via: ONR Code 458	1 Director
1	Chief, Bureau of Ships	U.S. Navy Underwater Sound
	Department of the Navy	Reference Laboratory
	Washington 25, D. C.	Box 8337
	Attn: Code 335	Orlando, Florida
	(1 ea. to:) Code 687	1 Director
	670	Ordnance Research Laboratory
	450	Pennsylvania State University
		P.O. Box 30
1	Chief of Naval Operations	University Park, Pennsylvania
	Department of the Navy	Via: Production Representa-
	Washington 25, D. C.	tive
	Attn: Pers-153	
1	Commander Submarine Force	1 Director
	U.S. Atlantic Fleet	University of Texas Defense
	Fleet Post Office	Research Laboratory
	New York, New York	P.O. Box 8029
		Austin, Texas
		Via: ONR Res. Rep.
1	Commander Submarine Force	University of Texas
	U.S. Pacific Fleet	
	Fleet Post Office	1 Commanding Officer and
	San Francisco, California	Director
1	Commander Submarine	Naval Ordnance Laboratory
	Flotilla ONE	White Oak
	Fleet Post Office	Silver Spring 19, Maryland
	San Francisco, California	Attn: RA Division
		1 Commander
		Submarine Development
		Group TWO
		Fleet Post Office
		New York, New York

CONFIDENTIAL

CONFIDENTIAL

- 1 Commander Destroyer
Development Group
U.S. Pacific Fleet
U.S. Naval Shipyard
Long Beach, California
- 1 Deputy Commander Pacific
Operational Test and
Evaluation Force
U.S. Naval Air Station
North Island
San Diego 35, California
- 1 Commanding Officer and
Director
U.S. Navy Engineering
Experiment Station
Annapolis, Maryland
- 1 Commanding Officer and
Director
U.S. Naval Ordnance
Test Station
China Lake, California
- 1 Institute for Defense
Analysis
Communications Research
Division
Von Neuman Hall
Princeton, New Jersey
- 1 Director
Woods Hole Oceanographic
Institution
Woods Hole, Massachusetts
- 1 Director
Marine Physical Laboratory
San Diego 52, California
- 1 Chief of Naval Operations
Department of the Navy
Washington 25, D. C.
Attn: Op 07

CONFIDENTIAL